Exhibit 2

Chapters from
Current Protocols in Molecular Biology
Volume 4

CURRENT PROTOCOLS IN MOLECULAR BIOLOGY

VOLUME 4

5

CONSTRUCTION OF RECOMBINANT DNA LIBRARIES

continued

CHAPTER 5
Construction of Recombinant DNA Libraries

INTRODUCTION

Construction of recombinant DNA molecules by simply ligating vector DNA and a fragment of interest is a straightforward process, discussed in Chapters 1 and 3. Special problems arise, however, when the fragment of interest represents only a very small fraction of the total target DNA. This is the case in two commonly encountered situations: isolation of single copy genes from a complex genome and isolation of rare cDNA clones derived from a complex mRNA population. This chapter describes techniques to generate recombinant DNA libraries which contain complete representation of genomic or cDNA sequences. Chapter 6 describes strategies and protocols for isolating particular desired sequences from such libraries.

The DNA of higher organisms is remarkably complex: a mammalian haploid genome contains approximately 3×10^9 base pairs. A particular 3000-bp fragment of interest thus comprises only 1 part in 10^6 of a preparation of genomic DNA. Similarly, a particularly rare mRNA species may comprise only 1 part in 10^5 or 10^6 of total poly(A) containing

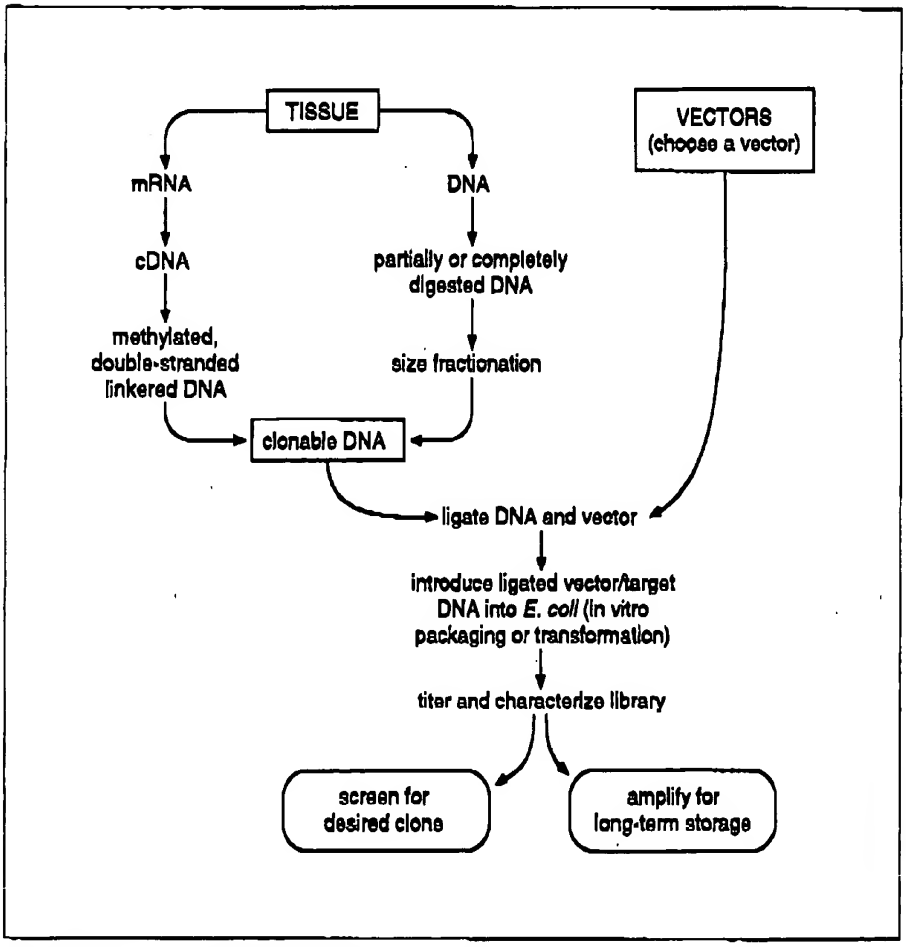

Figure 5.0.1 Steps involved in the construction of cDNA or genomic DNA libraries.

RNA, a ratio that is usually unaffected by the process of copying the RNA into cDNA. Clearly, the main problem in generating a useful recombinant DNA library from either genomic DNA or cDNA is the creation of the huge population of clones necessary to ensure that the library contains at least one version of every sequence of interest. The solutions to this problem are basically similar for genomic and cDNA libraries. As diagrammed in Fig. 5.0.1, the genomic DNA or cDNA are first prepared for insertion into the chosen vector. The vector and target DNA are then ligated together and introduced into *E. coli* by either packaging into phage λ heads in vitro or by direct transformation. In some aspects, however, strategies for isolation of individual genomic or cDNA clones can be quite different. The particular problems of creating these two different types of libraries will be discussed in detail separately (*UNITS 5.1 & 5.2*).

Most cDNA library screening procedures involve positive identification of cDNA clones with either antibodies or hybridizing nucleic acid probes. Subtracted cDNA libraries provide a method for identifying mRNAs (as cDNAs) that are expressed in one cell but not another. cDNAs are synthesized from the cell expressing the desired mRNA and all of the sequences expressed in a cell not expressing the RNA are removed by hybridization and selection. The remaining sequences are cloned into a bacteriophage or plasmid vector to produce a subtracted cDNA library. Estimation of the number of clones that must be screened is difficult, as it will vary with the cell type and the gene to be identified.

The *E. coli* vectors described in this chapter are limited with regard to the size of insert DNA that can be accommodated (~20 kb for lambda and ~40 kb for cosmid vectors). The ability to clone much larger fragments of DNA, however, has become an essential requirement for many genome analysis projects. Yeast artificial chromosome (YAC) vectors, maintained in yeast hosts, typically carry inserts ranging from 0.3 to 1.2 Mb of genomic DNA. Both the size and complexity of YAC libraries pose special considerations for production, screening, and analysis, and these concerns are addressed in *UNITS 6.9 & 6.10* in the following chapter.

Two important general points pertain to both genomic and cDNA libraries. First, it is essential that both the vector DNA and target DNA used to create the library are not contaminated by exogenous sequences detectable by the probes that will be used to isolate the clones of interest. There are obvious, potentially disastrous effects of contaminated target DNA—for example, by only 1 part in 10^5 of a plasmid containing the cDNA sequences to be used as a probe. See "Going for the gene," *The Boston Globe Magazine*, Aug. 2, 1987, for an account of such a mistake. Common sense dictates care and use of absolutely clean and, where possible, disposable materials throughout.

Second, by far the simplest way to generate a library is to "clone by phone" and get one from somebody else. Many useful libraries, including examples of human and other mammalian genomic or cDNA libraries, have been made over the years, and investigators are frequently willing to send out libraries which they have created. In some cases, journals (e.g., *Cell, Science, Proceedings of the National Academy of Sciences U.S.A.*, and the publications of the American Society for Microbiology) require that libraries and individual clones discussed in their pages be freely available to other investigators. In addition, YAC libraries are produced, maintained, and generally available from large "core" academic or institutional laboratories at Washington University (St. Louis), the ICRF (London), and CEPH (Paris); see *UNIT 6.9* for details. Both stock and custom-made libraries are also available from a variety of commercial sources.

J.G. Seidman

OVERVIEW OF RECOMBINANT DNA LIBRARIES

The units in this section present strategies for generating genomic DNA and cDNA libraries (*UNITS 5.1 & 5.2*, respectively). Production of recombinant DNA libraries can be a very laborious procedure. We strongly recommend the purchase of genomic or cDNA libraries when possible; otherwise, we recommend purchasing reagent kits for producing libraries. These kits can save considerable time and effort.

Genomic DNA Libraries

Genomic DNA libraries are almost always screened by hybridization using a radioactive nucleic acid probe. Since this approach is essentially independent of a particular vector or type of target DNA, the main problem faced when considering creation of a genomic DNA library is simply generating a large enough number of recombinant DNA clones. The basic strategies used to address this problem have included both minimizing the number of clones necessary by incorporating large fragments of genomic DNA, and maximizing cloning efficiency by using vectors based on bacteriophage λ. This unit will discuss the appropriate numerical considerations for both ordinary genomic DNA libraries and subgenomic DNA libraries, and will then describe a limited number of appropriate vectors.

REPRESENTATION AND RANDOMNESS

The size of a library of completely random fragments of genomic DNA necessary to ensure representation of a particular sequence of interest is dictated by the size of the cloned fragments and the size of the genome. The likelihood that a sequence of interest is present in such a random library can be estimated by simple statistics based on the Poisson distribution (Clarke and Carbon, 1976). Specifically, the number of independent clones, N, that must be screened to isolate a particular sequence with probability P is given by

$$N = \ln(1 - P)/\ln[1 - (I/G)]$$

where I is the size of the average cloned fragment, in base pairs, and G is the size of the target genome, in base pairs.

For a 99% chance of isolating an individual sequence from a typical mammalian genome using a typical phage λ vector

$$N = \ln(1 - 0.99)/\ln[1 - (2 \times 10^4/3 \times 10^9)] = 690{,}000.$$

This equation can be used to define a useful rule-of-thumb by calculating the probability of isolating a fragment of interest as a function of (I/G). In general, to have a 99% chance of isolating a desired sequence, the number of clones screened should be such that the total number of base pairs present in the clones screened ($I \times N$) represents a 4.6-fold excess over the total number of base pairs in the genome (G) (Seed et al., 1982).

When the desired fragment can be purified, the size of the library can then be estimated by

$$N \cong 3 \times 1/p$$

where p = the probability of isolating a particular fragment = 1/total number of fragments in the pool.

It is important to note that this simple analysis assumes that the cloned DNA segments randomly represent the sequences present in the genome. This assumption is true only if the target DNA was cleaved completely randomly prior to insertion into the vector. In the strictest sense, this level of randomness can be approached only by the relatively inconvenient means of shearing the target DNA.

With common sense and care, however, sufficiently random cleavage of target DNA can generally be obtained using partial digestions with restriction enzymes (*UNIT 5.3*). One simple limitation of this approach is that fragments which are larger than the capacity of the vector as complete digestion products will be excluded from the library. Clearly, it is best to use an enzyme that cuts the DNA of interest both frequently and without any bias in selection of one site over another (such bias is seen with *Eco*RI, for example). The enzyme *Sau*3A, which recognizes the 4-bp site GATC and generates fragments compatible with several convenient phage λ and cosmid vectors (see below), has proved useful for generating partial digestion libraries.

Given an enzyme that cleaves frequently and randomly, it is not so obvious which partial digestion protocol will lead to the most random products. As described in detail by Seed et al. (1982), randomness is maximized by partial digestion to an extent where the size of the numerically most abundant class of partial products equals the vector capacity. This is not reflected by an extent of digestion in which the location on a gel of the maximum staining intensity of partial digestion products equals the fragment size of interest (a consequence of the fact that larger fragments stain more intensely than an equal number of smaller fragments). To determine the optimal extent of digestion, resolve a series of partial digestions of increasing extent on a gel and examine the amount of staining in *only* the size class of interest. The lane in which the greatest amount of DNA is seen corresponds to *twice* the appropriate extent of digestion (Seed et al., 1982).

SUBGENOMIC DNA LIBRARIES

Sometimes only a small and relatively well characterized fragment is desired. For example, if a particular 1-kb *Bam*H1 fragment is of interest, it can be purified and used to generate a smaller, potentially easier to screen library. Such libraries, which represent only a fraction of the genome, are called subgenomic DNA libraries.

Numerical considerations show that maximizing the fold of purification of target DNA is crucial for subgenomic DNA libraries. One can use the equation described above to estimate the number of clones necessary by simply assuming that the genome size is reduced by the amount of purification. For example, if the desired 1-kb mammalian DNA fragment was purified 10-fold from the rest of the genomic DNA, and the resultant library was otherwise a random representation of the remaining 10% of the genome, then

$$N = \ln(1 - 0.99)/\ln[1 - (1 \times 10^3/3 \times 10^8)] = 1{,}380{,}000.$$

In this case the subgenomic approach has actually *increased* the number of clones necessary as compared to the random library, due to the large decrease in the size of the insert. Increasing purification by another factor of 10 decreases the number of clones needed by a factor of 10. As a minimum, the fold of purification must exceed the ratio of genomic DNA library insert size to subgenomic insert size.

A simple way to increase the fold of purification is to use multiple, sequential digestion strategies in cases where details of the restriction map of the sequences of interest are known. After initial purification of a given fragment, redigestion with another enzyme that gives a smaller (clonable) fragment will generally yield significant further purification relative to the original DNA.

VECTORS FOR GENOMIC DNA LIBRARIES

Because of their combination of high cloning efficiency and relatively large insert size, either bacteriophage λ vectors or hybrid plasmid vectors called cosmids (which contain particular λ sequences that direct insertion of DNA into phage particles) are generally used to construct genomic DNA libraries. The biology and general properties of these two types of vectors are described in *UNIT 1.10*.

Briefly, the high cloning efficiency of both types of vectors is a consequence of the ability of simple extracts of phage λ-infected cells to insert exogenously added λ DNA, or recombinant DNA containing appropriate λ sequences, into preformed λ heads and tails, generating infectious phage particles. Up to 10% of added concatameric phage genomes can be packaged in this way, an efficiency significantly greater than that of introduced plasmid DNA into *E. coli* by transformation. In the case of cosmid vectors, the recombinant DNA inserted into the phage contains plasmid vector sequences and replicates as a plasmid after infecting bacteria (see *UNIT 1.10*).

The cloning capacity of both types of vectors is dependent on the size of DNA that can be accommodated by λ phage heads, approximately 35 to 50 kb. However, the vectors differ significantly in the fraction of this total taken up by vector sequences. Most phage vectors designed for genomic DNA libraries can accommodate foreign DNA fragments of 10 to 20 kb generated by a limited variety of restriction enzymes. Cosmids can generally accept 30- to 40-kb fragments generated by any of a number of restriction enzymes.

The choice between phage and cosmid vectors is generally based on the size of the desired genomic DNA segment. Most investigators feel that phage libraries are easier to handle, and choose a phage vector if the desired segment is less than ~20 kb. Larger segments require the use of cosmid vectors.

Bacteriophage λ Vectors

Significant design advances over the years have resulted in the development of several

easy-to-use phage λ vectors. These vectors have two basic features in common: ability to accept fragments generated by several restriction enzymes, and biochemical and/or genetic selection against the so-called stuffer sequences present in the original vector in the place of the exogenously added DNA. As described in *UNIT 1.10*, this stuffer fragment is necessary because vectors that contain the minimum segment of the λ genome necessary for phage propagation (approximately 30 kb) are too small to be packaged into normal phage heads. Earlier vectors required rather laborious biochemical separations of vector and stuffer fragments prior to insertion of foreign DNA.

The minimal λ genome contains restriction sites for a number of enzymes frequently used for cloning in plasmid vectors. Newer vectors have partially circumvented this problem by eliminating some of these sites (notably those for *Bam*HI) and addition of new, unique sites to polylinkers which flank the stuffer fragment.

The segment of the λ genome that can be replaced by exogenous DNA contains genes whose expression prevents phage growth in bacterial hosts containing P2 prophages (see *UNIT 1.10*). Engineering the stuffer fragment to express these genes regardless of orientation relative to the rest of the genome has resulted in vectors that will grow on such P2 lysogens only if they have incorporated foreign DNA in the place of the stuffer. Such vectors obviate the need to physically remove the stuffer prior to ligation to exogenous DNA.

An alternative biochemical strategy to prevent reinsertion of the stuffer segment has also been developed. In vectors designed for this approach, two identical polylinkers flank the stuffer in an inverted orientation relative to each other. In this arrangement, double digestion of the polylinkers with two appropriate enzymes generates vector and stuffer fragments with heterologous, nonligatable ends. Simple, preferential ethanol precipitations (see *UNIT 2.1*) remove the very small polylinker fragments. Since the vector segment by itself is too short to generate a viable phage, all plaques generated by in vitro packaging of a ligation of vector and insert DNA should be recombinant.

The vector λEMBL3 allows both the genetic and the biochemical strategies to avoid purification of stuffer fragments, and includes several useful cloning sites in the polylinker. This versatile and modern vector has been successfully used to create many libraries.

Cosmid Vectors

Any plasmid cloning vector that contains the λ *cos* site can be used as a cosmid. A number of cosmid vectors designed for particular applications include additional elements such as genes that allow transfer to non-*E. coli* bacterial cells or dominant markers for selection in mammalian cells. Such add-ons decrease the cloning capacity of the vector and should be avoided if possible. One useful, simple cosmid vector is pJB8 (Ish-Horowitz and Burke, 1981), a 5.4-kb plasmid that accepts genomic DNA digested with *Sau*3A and can be used with several cosmid cloning strategies.

Vectors for Subgenomic DNA Libraries

It is possible to use simple plasmid vectors for subgenomic DNA libraries if the level of purification and recovery of the target fragment is sufficient to overcome the relative inefficiency inherent in plasmid cloning. In general, however, phage λ vectors designed for direct insertion of foreign DNA rather than substitution for a stuffer fragment are used. A very large number of potential insertion vectors exist to accommodate fragments generated by a variety of restriction fragments: wild-type λ is a natural insertion vector which should accommodate *Sal*I or *Xho*I fragments up to 2 to 3 kb. λgt10 (see *UNIT 1.10*) is the only vector in general use which allows selection against nonrecombinant phages, and is recommended for cloning *Eco*RI fragments.

LITERATURE CITED

Clark, L. and Carbon, J. 1976. A colony bank containing synthetic *ColE1* hybrids representative of the entire *E. coli* genome. *Cell* 9:91-99.

Ish-Horowitz, D. and Burke, J.F. 1981. Rapid and efficient cosmid vector cloning. *Nucl. Acids Res.* 9:2989-2999.

Seed, B., Parker, R.C., and Davidson, N. 1982. Representation of DNA sequences in recombinant DNA libraries prepared by restriction enzyme partial digestion. *Gene* 19:201-209.

Contributed by David D. Moore
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

UNIT 5.2

cDNA Libraries

The most basic step in constructing a cDNA library is the process of generating a double-stranded DNA copy of the mRNA. In the last few years, preparation of this cDNA has been simplified by improved strategies and availability of higher quality enzymes. Thus, it should be straightforward to obtain essentially full-length cDNA copies for mRNAs up to the 3 to 4 kb range, and at least feasible for even larger mRNAs. As described in detail in the cDNA protocols (*UNITS 5.5 & UNIT 5.6*), the most important factor affecting quality of cDNA is the quality of the mRNA. Particularly for a large message, it is essential to start with the highest quality RNA available.

Two related issues dominate the strategies for constructing cDNA libraries. The first is the relative abundance of the clone of interest, which can vary over a wide range. Highly abundant messages can represent 10% or more the total mRNA, whereas very rare messages can be as low as 1 part in 10^6, particularly if the gene of interest is only being expressed in a fraction of the cells used as a source of mRNA. The second issue is the screening method (see Chapter 6), which can range from simply sequencing several individual isolates until the desired clone is identified, through ordinary hybridization methods, to complex strategies involving expression of identifiable antigens or biological activities.

Obviously, the size of the library necessary to include the clone of interest is a direct reflection of the relative abundance of the mRNA of interest. In general, however, this abundance is not known with precision. In addition, the representation of some sequences in the cDNA library, particularly the 5′ ends of large messages, will be less than expected from their mRNA abundance. It is sensible to aim for a library that contains at least 5 times more recombinants than the total indicated by the lowest abundance estimate. In some cases this number should be multiplied by various factors based on screening efficiency. If it is necessary to fuse a peptide-coding region to a vector in a particular reading frame, for example, the number of identifiable clones is only ⅙ of those present in the library.

If the mRNA of interest is relatively abundant, efficiency of generating clones is not so important, and the choice of cloning strategy and vector should be based on the desired use for the clone. If, for example, expression in *E. coli* is the object, the cDNA library can be inserted directly into an appropriate expression vector. This might involve choosing linkers or adaptors useful for insertion into the vector, and simple screening by hybridization.

In many cases, however, the mRNA of interest is relatively rare, and high cloning efficiency is of central importance. As with genomic libraries, this has led to development and use of phage λ vectors. In general, there are two types of λ vectors for cDNA cloning adapted for the two most common methods of library screening.

If the library is to be screened by hybridization with a nucleic acid probe (*UNITS 6.1 & 6.3* or *6.4*), any insertion vector is appropriate. A vector that is particularly good for this approach is λgt10. As mentioned above, this insertion vector allows direct selection against nonrecombinant phages. A useful feature of this vector for cDNA cloning is that it accepts *Eco*RI inserts. Methylation of the double-stranded cDNA with *Eco*RI methylase and addition of *Eco*RI linkers is an efficient way to generate clonable cDNA (see *UNITS 5.5 & 5.6*).

If the library is to be screened by use of antibody probes (*UNIT 6.7*), it is necessary to use an appropriate *E. coli* expression vector. In general, such vectors are based on expression of a fusion protein in which a segment of the peptide of interest is fused to a highly expressed, stable *E. coli* protein. The most commonly used expression vector is λgt11 (*UNIT 1.10*), in which the cloned peptide coding sequences are fused to coding sequences for β-galactosidase.

KEY REFERENCES

Frischauf, A.-M., Lehrach, H., Poustka, A., and Murray, N. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* 170:827.

Huynh, T., Young, R., and Davis, R. 1984. Construction and screening cDNA libraries in λgt10 and λgt11. *In* DNA Cloning, Vol. 1: A Practical Approach (D. Glover, ed.) pp. 49-78. IRL Press, Oxford.

Contributed by David D. Moore
Massachusetts General Hospital
and Harvard Medical School
Boston, Massachusetts

PREPARATION OF INSERT DNA FROM GENOMIC DNA

Completely or partially digested genomic DNA must be size fractionated before ligation to vector to remove irrelevant small and large fragments. If size fractionation is not performed, small fragments will ligate together and produce recombinants that are difficult to analyze. Large insert DNA fragments will not allow the vector to grow, but will ligate to vector DNA and alter vector DNA requirements. Size-purified fragments isolated from complete digests of genomic DNA are less complex than the entire genome and thus reduce the number of DNA clones that must be produced in order to obtain the desired subgenomic library.

Procedures for making insert fragments for genomic and subgenomic DNA libraries involve digestion of DNA followed by size fractionation. DNA is either partially digested for preparation of complete genomic libraries or completely digested for preparation of subgenomic libraries. Basic protocols for digesting DNA with restriction enzymes are presented in *UNIT 3.1*; however, *UNIT 5.3* presents support protocols with necessary modifications to ensure that large amounts of genomic DNA are properly digested.

Two methods for size fractionation of genomic DNA are presented. Both protocols are appropriate for the isolation of DNA fragments that will subsequently be used for library construction. These protocols aim to maximize DNA fragment yield while minimizing exposure of the DNA to reagents or conditions that inhibit subsequent ligation to vector and introduction of hybrid molecules into the bacterial cell. The sucrose gradient and preparative gel electrophoresis methods circumvent problems found in other fractionation protocols.

Sucrose gradient fractionation is generally faster than preparative gel electrophoresis. However, the latter procedure has a higher capacity for resolving large amounts of DNA, is applicable to a larger range of sizes, and has significantly better resolution.

These procedures emphasize the requirements for fractionating large quantities of DNA and producing DNA that will ligate to vector. Normally, large quantities of genomic DNA can be obtained from the species of interest so that when producing genomic libraries the investigator has the luxury of being able to work with more DNA than when producing cDNA libraries. The procedures used here all assume that genomic DNA is available (see *UNITS 2.1 & 2.4* for DNA prep procedures). The amount of DNA required varies depending on the complexity of the genome being used.

NOTE: *For laboratories using recombinant DNA techniques and isolating large quantities of plasmid, bacteriophage, or cosmid DNA, remember that the smallest amount of contamination of genomic DNA with recombinant DNA is disastrous. Contamination of genomic DNA at 1 ppm with recombinant plasmid or bacteriophage will cause great difficulty because they may grow and be identified during screening procedures as the desired clones. Thus, all plasticware, glassware, and reagents used for the preparation of genomic DNA or mRNA and cDNA should be maintained separately from those used for plasmid or bacteriophage DNA preparation. The extensive use of disposable plasticware is strongly recommended.*

UNIT 5.3

BASIC PROTOCOL

Size Fractionation Using Sucrose Gradients

SUCROSE GRADIENT PREPARATION OF SIZE-SELECTED DNA

Partially or fully digested DNA consists of a population of DNA fragments ranging in size from hundreds of base pairs to over 100,000 bp in length. This protocol effectively separates such a mixture of DNA fragments into different size classes. To accomplish this the DNA solution is heated to dissociate aggregated DNA fragments and is then loaded onto a high-salt sucrose gradient. After centrifugation and gradient fractionation, the appropriate fractions are identified by agarose gel electrophoresis. This protocol can also be used to purify bacteriophage λ vector arms.

Materials

Completely or partially digested genomic DNA (support protocols)
STE buffer (*APPENDIX 2*)
10% and 40% sucrose solution
0.9% agarose gel
100% ethanol
TE buffer (*APPENDIX 2*)
Sucrose gradient maker
Beckman SW-28 or SW-41 rotor or equivalent
Additional reagents and equipment for ethanol precipitation (*UNIT 2.1*) and agarose gel electrophoresis (*UNIT 2.5*)

1. Begin with partially or fully restriction-enzyme digested DNA (see support protocol) at a concentration of about 1 mg/ml in STE buffer.

   It is essential that the DNA be completely dissolved.

2. Prepare a linear 10% to 40% sucrose gradient in an SW-28 centrifuge tube (38-ml gradient) or SW-41 tube (12-ml gradient).

   There are a variety of gradient makers; follow the manufacturer's instructions for preparing the gradient.

3. While the gradient is being poured, heat the digested DNA to 65°C for 5 min to dissociate any DNA aggregates.

4. Carefully layer the DNA solution on top of the sucrose gradient.

   Do not exceed 0.5 mg genomic DNA per SW-28 centrifuge tube or 0.2 mg genomic DNA per SW-41 tube. For λ vector DNA, do not exceed 50 μg and 20 μg, respectively (see critical parameters).

5. Centrifuge at 20°C, 113,000 × *g* (25,000 rpm in SW-28 rotor) or 154,000 × *g* (30,000 rpm in SW-41 rotor) for 16 to 24 hr, depending upon the size of the desired DNA fragments. For cosmid-sized inserts (40,000 bp) centrifuge 16 to 18 hr. For phage-sized inserts (18,000 bp) centrifuge 24 hr.

6. Fractionate the gradient by carefully placing a capillary tube at the bottom of the centrifuge tube and pumping out the gradient, heavier fractions first. This prevents the mixing of smaller DNA fragments with the desired larger fragments. Collect 750-μl fractions into microcentrifuge tubes (see Fig. 5.3.1A).

   Note: 12-ml gradients can also be fractionated by removing 750-μl aliquots from the top with a mechanical pipettor (e.g., P1000 Eppendorf pipet).

7. Determine the size of the collected DNA fractions by electrophoresing 40-μl samples of the gradient in a 0.9% agarose gel at high voltage (see *UNIT 2.5*).

SCREENING OF RECOMBINANT DNA LIBRARIES

6

continued

V YEAST ARTIFICIAL CHROMOSOME LIBRARIES

VI SPECIALIZED STRATEGIES FOR SCREENING LIBRARIES

CHAPTER 6
Screening of Recombinant DNA Libraries

INTRODUCTION

The usual approach to isolating a recombinant DNA clone encoding a particular gene or mRNA sequence is to screen a recombinant DNA library. As described in Chapter 5, a recombinant DNA library consists of a large number of recombinant DNA clones, each one of which contains a different segment of foreign DNA. Since only a few of the thousands of clones in the library encode the desired nucleic acid sequence, the investigator must devise a procedure for identifying the desired clones. The optimal procedure for isolating the desired clone involves a positive selection for a particular nucleic acid sequence. If the desired gene confers a phenotype that can be selected in bacteria, then the desired clone can be isolated under selective conditions (*UNIT 1.4*). However, most eukaryotic genes and even many bacterial sequences do not encode a gene with a selectable function. Clones encoding nonselectable sequences are identified by screening libraries: the desired clone is identified either because (1) it hybridizes to a nucleic acid probe, (2) it expresses a segment of protein that can be recognized by an antibody, or (3) it promotes amplification of a sequence defined by a particular set of primers.

Screening libraries involves the development of a rapid assay to determine whether a particular clone contains the desired nucleic acid sequence. This assay is used first to identify the recombinant DNA clone in the library and then to purify the clone (see Fig. 6.0.1). Normally, this screening procedure is performed on bacterial colonies containing plasmids or cosmids or on bacteriophage plaques. To test a large number of clones at one time, the library is spread out on agarose plates (*UNIT 6.1*), then the clones are transferred to filter membranes (*UNIT 6.2*). The clones can be simultaneously hybridized to a particular probe (*UNITS 6.3 & 6.4*) or bound to an antibody (*UNITS 6.7 & 6.11*). When the desired clone is first identified, it is usually found among many undesirable clones; an important feature

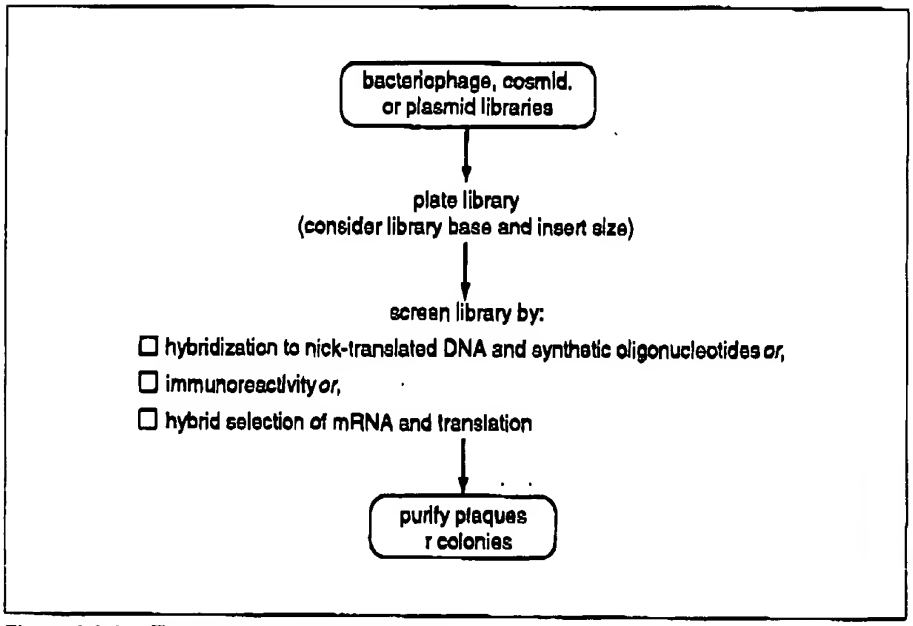

Figure 6.0.1 Flow chart for screening libraries.

of library screening is the isolation of the desired clones (*UNITS 6.5, 6.6 & 6.12*). Another method for identifying the desired clone involves hybrid selection (*UNIT 6.8*), a procedure by which the clone is used to select its mRNA. This mRNA is characterized by its translation into the desired protein. Libraries consisting of large genomic DNA fragments (~1 Mb) carried in yeast artificial chromosome (YAC) vectors have proven to be tremendously useful for genome analysis. In general, these libraries (which are usually produced by large "core" laboratories) are intially screened using a locus-specific PCR assay (*UNIT 6.9*); the clone resulting from the initial round of screening is subsequently analyzed by more conventional hybridization methods (*UNIT 6.10*).

To screen a DNA library, one must first devise the screening procedure. The next important choice is the selection of a recombinant DNA library. When choosing which library to screen the investigator should consider whether he or she wants to isolate clones encoding the gene or the mRNA sequence. cDNA clones encode the mRNA sequence and allow prediction of the amino acid sequence, whereas genomic clones may contain regulatory as well as coding (exon) and noncoding (intron) sequences. The differences between genomic and cDNA libraries are discussed in Chapter 5.

Another critical parameter to be determined before proceeding with a library screen is the number of clones in the library that must be screened in order to identify the desired clone. That is, what is the frequency of the desired clone in the library? This frequency is predicted differently for genomic and cDNA libraries, as described below.

Screening a genomic library. In general, genomic libraries can be made from DNA derived from any tissue, because only two copies of the gene are present per cell or per diploid genome. The predicted frequency of any particular sequence should be identical to the predicted frequency for any other sequence in the same genome. The formula for predicting the number of clones that must be screened to have a given probability of success is presented in *UNIT 5.1*. This number is a function of the complexity of the genome and the average size of the inserts in the library clones. For amplified libraries, the base (see *UNIT 5.1*) must exceed this number. Usually about 1 million bacteriophage clones or 500,000 cosmid clones must be screened to identify a genomic clone from a mammalian DNA library. Many of the clones that are screened from an amplified library will be screened more than once; the total number of clones that must be screened is 30 to 40% greater than the number calculated by the formula.

Screening a cDNA library. The optimal cDNA library is one made from a particular tissue or cell that expresses the desired mRNA sequence at high levels. In highly differentiated cells, a particular mRNA may comprise as many as 1 of 20 of the poly(A)$^+$ mRNA molecules, while some mRNAs are either not present at all or comprise as low as 1 molecule in 100,000 poly(A)$^+$ mRNA molecules. When choosing a cDNA library the investigator must make every effort to obtain a library from a cell where the mRNA is being expressed in large amounts. Of course, the number of clones that must be screened is determined by the abundance of the mRNA in the cell. The amount of protein that is found in the cell is frequently a good indicator of the abundance of the mRNA. Thus, proteins that comprise 1% of the total cell protein are made by mRNAs that usually comprise 1% of the total poly(A)$^+$ mRNA, and the desired cDNA clones should comprise about 1% of the clones in the cDNA library.

Screening a YAC library. In the typical genomic libraries maintained in *E. coli* (described in Chapter 5), the size of the insert is limited to 20 to 25 kb for lambda vectors or to 40 to 45 kb for cosmid vectors. Yeast artifical chromosome (YAC) vectors, by contrast, are designed to carry much larger genomic DNA fragments and thereby facilitate genomic analysis, with inserts ranging from 0.3 to ~1 Mb in size. Conventional screening of YAC

libraries by hybridization is difficult, both because of the unfavorable signal-to-noise ratio and the sheer numbers of replica films required to represent an entire library.

For example, a standard YAC library representing 5 to 8 genome equivalents requires over 500 microtiter plates (and corresponding filters for screening by hybridization). Thus, most core laboratories screen YAC libraries using a locus-specific PCR assay whose primers define a particular sequence. The PCR screening is initially performed using pools (representing up to 4 microtiter plates or 384 YAC clones) or superpools (representing up to 20 microtiter plates or nearly 2000 clones), followed by subsequent rounds of screening to narrow down the possible candidates.

Specialized screening strategies. For particular applications, there exist specialized approaches to screening. For example, cloned cDNAs encoding cell surface or intracellular proteins can be identified by expression screening, involving rounds of transient expression of a library and subsequent screening by immunoselection (*UNIT 6.11*). The technique of recombination-based screening provides a rapid and efficient approach for screening a complex genomic library in bacteriophage lambda (*UNIT 6.12*). The library is screened for homology against a plasmid carrying a particular cloned target sequence. If homology exists, a recombination event occurs, resulting in integration of the plasmid into the phage, and the recombinant is isolated by genetic selection.

General considerations. When selecting the library it is critical that the base be larger than the number of clones to be screened. One problem with predicting the number of clones to screen is that most libraries are amplified and in the process of amplifying the library some clones are lost while others may grow more rapidly. Thus, if the desired clone is not found in a particular library, another independent library should be screened.

Having selected the library, the investigator is ready to begin screening for the desired clone. The technologies used to screen libraries are mostly extensions of the techniques that have been described earlier in the manual. Libraries are plated out, transferred to nitrocellulose filters, and hybridized to ^{32}P-labeled probes or bound to antibodies. The major problem associated with this technique is that "false" positives can be identified: the probe may hybridize to clones that do not encode the desired sequence. Approaches to minimize this problem are discussed in *UNIT 6.7*. A second source of undesired clones arises from the power of the screening procedures that are normally used to screen these libraries. The investigator will be screening as many as one million clones. If the library contains any contaminating recombinant DNA clones that have been previously grown in the laboratory, it will be identified in the screening procedure. Thus, extreme care must be exercised to prevent contamination of the library with previously isolated recombinant clones. Despite these problems the ability to screen large DNA libraries to isolate the desired clone provides a powerful tool for molecular biologists.

J.G. Seidman

PLATING LIBRARIES AND TRANSFER TO FILTER MEMBRANES

SECTION I

The basic principle of screening recombinant DNA libraries is that bacteriophage plaques, or bacterial colonies containing plasmids or cosmids, contain relatively large amounts of insert DNA that can be detected either directly by hybridization (see below) or indirectly by the protein that may be expressed from the cloned segment (*UNIT 6.7*). The first step in the nucleic acid hybridization screening procedure is to grow large numbers of colonies or plaques on agar plates. Replica copies of these colonies are transferred to nitrocellulose filters, where they can be screened. In this section the techniques for producing large numbers of colonies and plaques, and for transferring these to filter membranes, are discussed. Prerequisites to these procedures are that the library must already be chosen and the number of clones to be screened must be determined (see introduction to this chapter).

Plating and Transferring Bacteriophage Libraries

UNIT 6.1

BASIC PROTOCOL

Bacteriophage are plated onto agar plates at high density so that as many as 1 million different plaques can be screened. The bacteriophage plaques are then transferred to nitrocellulose filters, denatured, and baked. The library and the number of clones to be screened are predetermined. Principles for choosing the plaque density and the number of plates to be used are outlined in the commentary.

Materials

Host bacteria, selection strain if applicable (*UNIT 1.10*; Table 1.4.5; Table 5.10.1)
Recombinant phage (*UNIT 5.10*)
0.7% top agarose (prewarmed; *UNIT 1.1*)
82-mm *or* 150-mm LB plates; *or* 245 × 245–mm Nunc bioassay LB plates (*UNIT 1.1*)
0.2 M NaOH/1.5 M NaCl
0.4 M Tris·Cl, pH 7.6/2× SSC
2× SSC (*APPENDIX 2*)

Nitrocellulose membrane filters (or equivalent)
20-G needle
46 × 57–cm Whatman 3MM or equivalent filter paper
80°C vacuum oven *or* 42°C oven

Plating bacteriophage

1. Determine the titer of the library by serial dilution as described in *UNITS 1.11 & 5.7*.

   For λ vectors that allow genetic selection against nonrecombinants, plating should be done on the appropriate bacterial strain (e.g., P2 lysogen for EMBL vectors). LB plates should be poured several days in advance to allow them to dry prior to plating. The large Nunc plates are particularly prone to condensation on the surface of the agar, but this can be alleviated by allowing them to sit on the benchtop with covers removed for a few minutes to several hours before use.

2. Mix recombinant phage and plating bacteria (prepared as described in *UNIT 1.11*) in a culture tube as outlined in Table 6.1.1 and incubate 20 min at 37°C.

3. Add 0.7% top agarose to culture tube and transfer mixture to LB plates. Disperse bacteria and agarose on plates by tilting the plates back and forth. Mix cells and agarose for the large Nunc plates by gently inverting several times in a capped 50-ml tube prior to plating.

Top agarose rather than top agar should be used as agar tends to lift off with the nitrocellulose filter.

Melt the top agarose and cool to 45° to 50°C before use. If top agarose is too hot it will kill the bacteria, while if it is too cold the library will solidify in the tube.

4. Incubate plates at 37°C until plaques cover the plate but are not confluent. Incubation time varies between 6 and 12 hr and depends on type of phage and bacteria used. Store at 4°C.

Do not incubate unattended overnight, but rather place at 4°C and allow to continue growth the next day. Allowing phage plaques to incubate for the correct amount of time is critical. The object is to optimize two parameters. First, the plaques must be large enough to contain sufficient DNA to give a good signal. Second, if the plaques are too large and become confluent they are difficult to purify in subsequent steps. Because most nucleic acid probes give a very strong signal, we tend to prefer having smaller plaques and weaker signals.

5. Incubate plates at 4°C for at least 1 hr before applying filters.

Transferring to nitrocellulose filters

6. Label nitrocellulose filters with a ballpoint pen and apply face down (ink side up) on cold LB plates bearing bacteriophage plaques. This is best accomplished by touching first one edge of the filter to the agarose and progressively laying down more of the filter as it wets. Bubbles should be avoided. If difficulties are encountered the filter should not be adjusted on the plate, but rather removed and replaced with a new filter.

Nitrocellulose filters should be handled only with forceps or gloved hands.

7. Leave filters on plates for 1 to 10 min to allow transfer of phage particles to the filter. During this transfer period the orientation of the filter to the plate is recorded by stabbing a 20-G needle through the filter into the agar at several asymmetric points around the edge of the plate. Up to five replicas can be made from each plate. Remove the filter slowly from the plate with blunt, flat forceps and place face up on paper towels or filter paper.

Some investigators dip the needle used to orient the filter in India ink to more clearly mark the filter and agar. Other investigators mark the back of the agar plate with a black marker.

Making two replicas from each filter, hybridizing both to the DNA probe, and comparing the autoradiographs of the replica filters eliminates many possible artifacts. This is particularly helpful when screening with an oligonucleotide probe.

8. Dry the filters on the benchtop for at least 10 min.

This drying process binds the plaques to the filter.

Table 6.1.1 Recommended Mixtures for Plating Bacteriophage Libraries

LB plate ingredient	Plate size		
	82 mm	150 mm	245 × 245 mm[a]
Bacteria[b] (ml)	0.2	0.5	2
Phage, pfu	5,000	20,000-30,000	150,000
Top agarose, ml	3	7	30

[a]Nunc Bioassay plates distributed by Vangard International.

[b]Plating bacteria are prepared as described in Chapter 1.

Denaturation and baking

9. Place 46 × 57–mm Whatman 3MM paper on the benchtop and saturate with 0.2 M NaOH/1.5 M NaCl. Place filters on the paper face up for 1 to 2 min.

   The 3MM paper should be wet enough to allow immediate saturation of the filters, but not so wet that the solution pools on the surface.

10. Transfer filters (face up) to 3MM paper saturated with 0.4 M Tris·Cl, pH 7.6/2× SSC for 1 to 2 min and then to 3MM paper saturated with 2× SSC for 1 to 2 min.

    Some investigators immerse the filters in all three solutions. This procedure can make the plaques detected by hybridization appear diffuse.

11. Dry filters in a vacuum oven 90 to 120 min at 80°C or overnight in a regular oven at 42°C. Store at room temperature in folded paper towels or other absorbent paper until needed for hybridization (described in *UNIT 6.3* or *6.4*).

COMMENTARY

Background Information

There are two parts to this protocol—plating the library and preparing filters. The number of bacteriophage per plate determines the number of plates that must be poured. This number is defined by the number of recombinants in the library (i.e., base of the library) and the frequency of the expected clone in the library. There is no advantage to screening more than 3 to 5 times the base of the library. The frequency of the clone in the library is determined as follows.

cDNA libraries: the expected frequency of the desired RNA among the total RNA of the cell, ranging from $1/100$ to $1/50{,}000$.

Genomic libraries: the size of the insert divided by the total genome size.

Subgenomic libraries: the size of insert per total genome size times the fold purification of the DNA fragment (usually 10- to 50-fold).

The usefulness of a recombinant phage library depends on the ability to screen a large number of phage and identify the clone that carries the DNA sequence of interest. This has been made possible by the technique of in situ plaque hybridization described by Benton and Davis (1977). The phage are allowed to multiply in host bacteria in a thin layer of agarose on regular bacterial plates. When nitrocellulose is applied to the agarose, phage particles and unpackaged DNA adsorb to the filter to produce a replica of the plate surface. If the agarose surface is not excessively wet, there will be little spreading of the phage on the filter. Subsequent treatment of the filter with sodium hydroxide destroys the phage particles and denatures the phage DNA which then binds to the nitrocellulose. Neutralization of the filters is required to maintain the integrity of the nitrocellulose. Hybridization of these filters to a DNA or RNA probe will identify the location of the phage plaque of interest, which can then be recovered from the plate.

A common variation of this technique is the substitution of one of the nylon-based membranes for nitrocellulose (see *UNIT 2.9*). The advantage of nylon membranes is their durability, which allows multiple hybridizations to the same filter and allows one to sequentially clone several genes from the same library using a single set of filters. However, nylon filters do not offer an improvement in sensitivity and are often more expensive than nitrocellulose filter paper.

Literature Review

The molecular basis of λ phage replication and the adaptation of the λ genome for molecular cloning has been reviewed by Arber et al. (1983) and Williams and Blattner (1980). Principles governing the plating of λ phage have been outlined by Arber (1983); see also *UNIT 1.10*. Thorough understanding of these principles has led to a universal approach to plating phage libraries.

Critical Parameters

To prevent recombination between different phage, do not allow them to overgrow, and grow them in recombination-minus hosts where possible. Calculations of the amount of phage stock to be used per plate should be based on a recent titration, and plating cells should be fresh.

Filters must not become brittle during this procedure; brittle filters will be destroyed during the hybridization process. This can be

avoided by limiting the time in the hydroxide solution to less than 5 min, making certain that the 0.4 M Tris·Cl, pH 7.6/2× SSC brings the filters to neutral pH, and limiting the baking to 2 hr.

Troubleshooting

Plaques should be visible on the plate before filters are made. If there appears to be poor bacterial growth, it is possible that the top agarose was too warm and many bacteria were killed, or that the phage titer was higher than expected and most host cells were lysed. Lower than expected phage titer could be due to an inaccurate titration of the phage stock, poor host-cell preparation, or too little time for adsorption.

The preparation of the nitrocellulose filters will only be tested after hybridization is complete. Occasionally, hybridization to a plaque will produce a streak instead of a discrete circle on the autoradiograph, making location of the correct plaque difficult. Steps that will often correct this problem include: (1) drying plates with the cover removed for 1 to 2 hr before applying the filter, (2) drying the filters well before the hydroxide treatment, and (3) making certain that the face (phage side) of the filters is not directly in contact with the solutions.

Anticipated Results

This plating procedure characteristically produces plates with an even distribution of dense phage particles. It is sensitive enough to allow identification of a phage by hybridization even when the phage are plated at high density (>5000 plaques per 82-mm plate). A signal is easily visible after 18 to 24 hr, when filters are hybridized to a nick-translated DNA probe with activity of $>10^7$ counts/μg DNA.

Time Considerations

Usually plaques will become visible within 6 to 10 hr after plating. Bacteriophage should generally not be allowed to grow longer than necessary to visualize the plaques. Using the procedure outlined, even a large number of filters can be processed in a single day.

Literature Cited

Arber, W. 1983. A beginner's guide to lambda biology. *In* Lambda II (R.W. Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg, eds.) pp. 381-395. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Arber, W., Enquist, L., Hohn, B., Murray, N., and Murray, K. 1983. Experimental methods for use with lambda. *In* Lambda II (R.W., Hendrix, J.W. Roberts, F.W. Stahl, and R.A. Weisberg, eds.) pp. 433-466. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Williams, B.G. and Blattner, F.R. 1980. Bacteriophage lambda vectors for DNA cloning. *In* Genetic Engineering, Vol. 2 (J.K. Setlow and A. Mullander, eds.) p. 201. Plenum, NY.

Key References

Benton, W.D. and Davis, R.W. 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. *Science* 196:180-182.

Describes the method of plaque hybridization developed by the authors to allow isolation of phage possessing specific cloned DNA sequences.

Contributed by Thomas Quertermous
Massachusetts General Hospital
Boston, Massachusetts

Plating and Transferring Cosmid and Plasmid Libraries

BASIC PROTOCOL

A bacterial suspension is suctioned through a porous membrane, leaving the bacteria bound to the membrane surface. The membrane is transferred, bacteria up, to an agar plate upon which the bacteria will receive enough nutrients to grow into colonies. These filters can then be used for replica platings and for hybridization with specific DNA probes.

Materials

LB plates containing antibiotic (*UNIT 1.1*)
LB medium (*UNIT 1.1*)
LB plates containing 50 µg/ml chloramphenicol (*UNIT 1.1*)
0.5 M NaOH
1 M Tris·Cl, pH 7.5
0.5 M Tris·Cl, pH 7.5/1.25 M NaCl
10- or 15-cm Whatman 3MM or equivalent filter paper discs
Sintered glass filter with vacuum
Nitrocellulose membrane filters (10- or 15-cm, Millipore HATF)
20 × 20–cm Whatman 3MM or equivalent filter paper
20 × 20–cm glass plate
20-G needle
46 × 57–cm Whatman 3MM or equivalent filter paper
80°C vacuum oven

NOTE: All materials coming into contact with *E. coli* must be sterile.

Plating cosmids

1. Start with plasmid or cosmid library produced after transformation, transfection, or amplification (*UNIT 5.7*).

2. Determine titer of the library by serial dilutions using plates containing antibiotics (see *UNIT 1.3*).

   Remaining library suspension can be held at 4°C overnight with only minimal loss of viable bacteria.

   A 10-cm nitrocellulose filter can accommodate 10,000 to 20,000 colonies, while a 15-cm filter can hold up to 50,000.

3. Calculate the appropriate amount of the bacterial suspension for plating and dilute the suspension in LB medium such that there is the desired amount of bacteria in 5 ml (10-cm filter) or 10 ml (15-cm filter) of solution.

4. Meanwhile, prepare a layer of 10- or 15-cm Whatman 3MM paper discs on either the bottom part of a sintered glass Buchner funnel or on a porcelain filter funnel. Pour 10 to 20 ml LB medium over two or three layers of 3MM paper discs to make a level bed. The same pad of discs can be used for many filters.

   Sterilize filter apparatus and filter paper before use. The 3MM and nitrocellulose filters can be sterilized by autoclaving them while wrapped in aluminum foil.

   The purpose of this step is to spread the bacteria uniformly across the surface of a nitrocellulose filter. The filtering apparatus must be level, it must create a uniform suction to all the surface of the filter, and it should be easy to move the filters to and from the apparatus.

5. Label a nitrocellulose filter with a ballpoint pen on the side opposite that where the

bacteria will be plated. Place the filter on the surface of the LB/antibiotic plate to wet it.

The antibiotic plate must be permissive for cosmid- or plasmid-bearing bacterial cells and usually is ampicillin or tetracycline.

Most ballpoint pen inks do not smudge during the hybridization reaction. If the one you choose runs, try another type.

6. Remove the wet filter from an antibiotic plate to the filtration apparatus.

   The suction should be off.

   Carefully pipet the 5 to 10 ml of bacterial suspension onto the surface of the nitrocellulose filter, leaving the outer 4 to 5 mm of the filter free of solution.

   This outside bacteria-free ring leaves enough surface area to work with the filter without smearing or losing the colonies.

7. Slowly suction the solution down through the filter, taking care not to create any preferential suction pockets that would concentrate the bacteria. After suctioning all of the solution through the filter, transfer the filter back to the antibiotic plate on which it was wetted.

   In laying the filter down on the agar surface, take care to avoid trapping any air bubbles between the surface of the plate and the filter.

8. Plate the entire library in this way and incubate the plates upside down (agar side up) at 37°C until the colonies are ~1 mm in diameter.

   Do not overgrow the filters, as smaller colonies can be lost beneath larger, faster-growing recombinant bacteria.

Preparing replica filters

9. Label and wet another set of nitrocellulose filters, as described in step 5.

10. Remove the initial library filter from its plate and place on several sheets of 20 × 20 cm 3MM paper, bacteria side up. While wearing gloves, carefully position the wetted replica filter above the bacterial lawn. Lay the second filter upon the first, leaving the two filters offset by 2 to 3 mm.

    This overlap will help in the separation of the two filters after the replica transfer.

    Do not allow air bubbles to form between the two filters. These are excluded by touching the second filter to the first in the middle and then allowing the edges to fall.

11. Lay three sheets of 20 × 20–cm 3MM paper on the two filters, followed by a 20 × 20 cm glass plate. Using the palms of your hands, press with all your weight down on the glass plate, thus transferring the bacterial colonies from the library filter to the replica filter.

12. Remove the glass plate and the filter paper and, using a 20-G needle, punch holes 2 to 4 cm apart through both of the filters. These holes will allow the orientation of the film produced from the replica filter down on the library filter for the isolation of the correct clones.

13. Carefully peel the two filters apart, placing them both bacteria up, on their respective agar plates. Grow the replica colonies at 37°C overnight, leaving the library filters at 25°C overnight. After overnight growth, store the library filters on the agar plates at 4°C, while screening the replica filters.

    Multiple replica filters can be made from the same library filter. Incubate library filters 2 to 4 hr at 37°C or overnight at 25°C to allow regrowth of the colonies.

Then repeat steps 9 to 13. Normally, two copies of the cosmid are hybridized to each probe.

14. After the bacterial colonies have grown, the cosmids or plasmids on the replica filter are amplified by transferring them to an LB plate containing 50 μg/ml chloramphenicol and incubating at 37°C for 4 to 10 hr. This step will increase the signal produced by hybridization.

Preparing filters for hybridization

15. Remove the replica filters from the LB/chloramphenicol plates, place filters bacteria side up on a sheet of 46 × 57–cm 3MM paper soaked with 0.5 M NaOH, and leave them for 5 min.

16. Carefully transfer to a sheet of 46 × 57–cm 3MM paper soaked with 1 M Tris·Cl, pH 7.5. Allow neutralization to occur for 5 min.

17. Transfer to a third 46 × 57–cm filter soaked in 0.5 M Tris·Cl, pH 7.5/1.25 M NaCl. Neutralize 5 min.

18. Transfer filter to a dry sheet of 3MM paper to allow filter to dry.

After filters are completely dry, stack them on paper towels or other adsorbent paper. Each nitrocellulose filter should be separated by paper towels from other filters.

19. Transfer the stacked filters to a vacuum oven at 80°C for 90 min. Remove filters and hybridize with a nick-translated probe, as described in *UNITS 6.3* and *6.4*.

COMMENTARY

Background Information

There are two commonly used protocols for the screening of recombinant bacteria with hybridization probes. The first method involves the spreading of bacteria on the surface of agar using a sterile spreader (*UNIT 1.3*). A nitrocellulose membrane filter is then placed on top of the colonies and most of each colony is transferred to the filter. The filter is then treated as described in steps 15 to 19. This method works well when relatively small numbers of positive colonies are being selected (up to several thousand).

The second method employs a matrix of some type (here nitrocellulose filters are used) upon which bacteria can be plated and grown when the filter is placed on top of a nutrient agar surface. Once the plated bacteria have grown into visible colonies, the filters can be used for replica plating and in situ hybridization analysis.

Critical Parameters

In order to provide a uniform lawn of recombinant bacteria for screening, it is critical to ensure that the suction applied to the filters is uniform and not spotty. The best way to accomplish this is to suction the suspension through the filter slowly and to avoid any preferential suction sites in the filter. Make sure that the apparatus is level and that adequate layers of LB-soaked chromatography paper are used. Air bubbles will prevent bacterial growth, so be certain that air is not trapped between the filter and the agar surface.

Time Considerations

Once the apparatus is set up, it takes ~5 min per filter to wet the filter, suction the bacteria, and transfer to an LB plate. The colonies take ~15 hr to grow at 37°C, after which they can be transferred to 4°C until ready for the replica platings. Replica plating also requires 5 min per filter, and resulting filters will be ready for denaturation and hybridization after 15 hr at 37°C.

Key Reference

Hanahan, D. and Meselson, M. 1983. Plasmid screening at high density. *Meth. Enzymol.* 100:333-342.

Contributed by John H. Weis
Harvard Medical School
Boston, Massachusetts

SECTION II

HYBRIDIZATION WITH RADIOACTIVE PROBES

After plaques or colonies have been transferred to nitrocellulose filters, the desired clone can be detected by its ability to hybridize to a DNA probe. This is a rapid, effective screening procedure that allows the identification of a single clone within a population of millions of other clones. The filters are hybridized with a ^{32}P-labeled nucleic acid probe, the excess and incorrectly matched probe is washed off the filter, and the filter is autoradiographed. Two features of the nucleic acid probe used for these experiments are critical to the successful screening of recombinant DNA libraries. First, the probe must hybridize only to the desired clones and not to any other clones. Thus, the nucleic acid sequence used for a probe must not contain any reiterated sequences or sequences that will hybridize to the vector. Second, the specific activity of the probe must be at least 10^7 cpm/μg. Most of the procedures for labeling DNA or copy RNA molecules are described in Chapter 3, and a support protocol is presented here that allows the 5′ end-labeling of a mixture of oligonucleotides.

The two basic protocols presented in this section describe steps required to hybridize labeled probes to recombinant DNA clones on filters. Two protocols are presented because conditions for hybridizing short oligonucleotide probes and longer nucleic acid probes to filters are different.

UNIT 6.3

Using DNA Fragments as Probes

BASIC PROTOCOL

HYBRIDIZATION IN FORMAMIDE

Bacteriophage plaques or bacterial colonies bound to a filter membrane are detected by hybridization with a radioactive probe. Hybridization proceeds on prewet filters placed in a sealable plastic bag. After hybridization the filters are removed from the sealed bag, excess probe is washed off, and the filters are autoradiographed to identify the clones that have hybridized with the probe.

Materials

Nitrocellulose membrane filters bearing plaques, colonies, or DNA (*UNITS 6.1 & 6.2*)
Hybridization solution I
Radiolabeled probe, 1 to 15 ng/ml (*UNIT 3.5*)
2 mg/ml sonicated herring sperm DNA
High-stringency wash buffer I
Low-stringency wash buffer I
Sealable bags
42°C incubator
Water bath adjusted to washing temperature (see commentary)
Glass baking dish
Additional reagents and equipment for autoradiography (*APPENDIX 3*)

Incubate filters with probe

1. Wet filters with hybridization solution I. Lay a filter membrane bearing plaques on top of 5 to 20 ml of hybridization solution I and allow solution to seep through filter. It is important to wet only one surface at a time to prevent trapping air in filter. Wet each filter in turn, producing a stack of wet filters.

   When multiple filters are to be hybridized to the same probe, no more than twenty 8.2-cm discs or ten 20 × 20 cm square filters should be placed in one stack.

Estimate the volume of hybridization solution used to wet the filters; this is a significant fraction of the volume of the hybridization reaction.

2. Transfer the stack of wetted filters to an appropriately sized sealable bag. Add enough hybridization solution to generously cover filters and seal.

   Note the volume of hybridization solution used to cover the filters.

3. Prehybridize filters by placing the bag in a 42°C incubator for at least 1 hr.

   Some investigators omit this step.

4. While filters are prehybridizing, pipet the radioactive probe into a screw-cap tube, add 2 mg (1 ml) sonicated herring sperm DNA, and boil 10 min. Place boiled probes directly into ice to cool.

   The amount of probe used is important, and should be in the range of 1 to 15 ng/ml of hybridization reaction. The volume of the hybridization reaction can be assumed to be the amount of hybridization solution added to the filters.

5. Add 2 ml hybridization solution I to the boiled probe.

6. Remove bag containing filters from the 42°C incubator. Open bag, add probe mixture, exclude as many bubbles as possible, and reseal.

   A good way to add the radioactive probe is to take it up in a syringe with an 18-G needle and then inject it into the bag. Reseal the bag after adding probe.

7. Mix probe in the bag so that filter is evenly covered. Replace bag in the 42°C incubator and let hybridize overnight.

Wash filters to remove nonhybridized probe

8. Warm 1 liter high-stringency wash buffer I to the "washing temperature" in a water bath.

   The stability of washing temperature and salt concentrations are critical features of this experiment. See discussion in commentary.

9. Remove bag containing hybridizing filters from the 42°C incubator. Cut bag open and squeeze hybridization solution out of the bag.

   CAUTION: *Handle material carefully as it is extremely radioactive. This should be done on disposable paper bench covers.*

10. Quickly immerse the filters in 500 ml low-stringency wash buffer I at room temperature in a glass baking dish. Separate all the filters, as they may stick together during hybridization.

    The volume of the low-stringency wash buffer is not important as long as the filters are completely covered. The filters must not be allowed to dry as the radioactive probe will irreversibly bind the filters if the filters dry in contact with probe. (The type of container used to hold the filters is not important as long as it transfers heat well. Thus glass, metal, or enamel containers are better than plastic.)

    The low-stringency wash only removes nonhybridized probe formamide and hybridization solution; it does not determine the stringency of the hybridization.

11. Rinse the filters three times with 500 ml low-stringency wash buffer. Let the filters sit 10 to 15 min at room temperature in low-stringency wash buffer with each rinse.

12. Pour off the low-stringency wash buffer and pour in 500 ml high-stringency wash buffer (prewarmed to washing temperature).

13. Replace the high-stringency wash buffer with another 500 ml of high-stringency wash buffer, then place the glass dish containing the filters in incubator at wash temperature. Make sure that the temperature in the glass dish reaches the desired washing temperature by placing a thermometer directly into the bath and measuring the temperature. Usually 15 to 20 min at the desired wash temperature is sufficient to remove most of the background radioactivity.

Of course, if the glass dish is placed in a water bath, be careful that the water from the water bath does not get into the filters.

Autoradiographing filters

14. Remove filters and mount them either wet or dry on a plastic backing. If the filter(s) is to be exposed wet, then isolate it from the film by covering it with plastic wrap.

Used X-ray film provides a good form of plastic backing for filters.

15. Mark the filters with radioactive ink to assist in alignment and autoradiograph.

An easy way to apply radioactive ink is to mark adhesive-backed paper labels with radioactive ink and then attach the stickers to the plastic wrap cover.

X-ray intensifying screens greatly decrease the amount of exposure time required.

ALTERNATE PROTOCOL

HYBRIDIZATION IN AQUEOUS SOLUTION

This method differs mainly in that formamide is not used in the hybridization solution. Follow the basic protocol except use the reagents and alternate parameters listed below.

Additional Materials

Hybridization solution II
Low-stringency wash buffer II
High-stringency wash buffer II
65°C incubator

1. Prehybridize as in basic protocol except that the filters are prehybridized at 65°C using hybridization solution II.

Hybridization solution II may have to be prewarmed to solubilize the SDS.

2. Prepare probe as in step 4 of basic protocol and dilute with 2 ml of hybridization solution II.

3. Hybridize overnight as in steps 6 and 7 of basic protocol except use a hybridization temperature of 65°C.

4. Remove bag containing hybridization from the 65°C incubator. Squeeze out the hybridization solution, taking care to avoid contamination with the excess radioactive hybridization solution.

5. Immediately rinse filters twice with low-stringency wash buffer II.

It is unnecessary to maintain a given temperature for this wash; just let the filters sit in wash buffer at room temperature until ready to proceed.

6. At 65°C, proceed to wash filters with high-stringency wash buffer II. Employ multiple quick washes (5 to 8) and immerse filter in a final wash for ~20 min. Check the radioactivity of the filters with a Geiger counter and be certain that they produce a signal only a fewfold above background levels.

REAGENTS AND SOLUTIONS

High-stringency wash buffer I

0.2× SSC (*APPENDIX 2*)
0.1% sodium dodecyl sulfate (SDS)

High-stringency wash buffer II

1 mM Na_2EDTA
40 mM $NaHPO_4$, pH 7.2
1% SDS

Hybridization solution I

Mix following ingredients for range of volumes indicated (in milliliters):

Formamide	24	48	72	120	240	480
20× SSC	12	24	36	60	120	240
2 M Tris·Cl, pH 7.6	0.5	1.0	1.5	2.5	5.0	10
100× Denhardts solution	0.5	1.0	1.5	2.5	5.0	10
Deionized H_2O	2.5	5.0	7.5	12.5	25	50
50% dextran sulfate	10	20	30	50	100	200
10% SDS[a]	0.5	1	1.5	2.5	5	10
Total volume	50	100	150	250	500	1000

[a]In place of SDS, *N*-lauroylsarcosine (Sarkosyl) may be used.

Add the SDS last. The solution may be stored for prolonged periods at room temperature.

The dextran sulfate should be of high quality. Pharmacia produces acceptable-grade dextran sulfate. Recipes for SSC and Denhardt's solution are in APPENDIX 2.

Hybridization solution II

1% crystalline BSA (fraction V)
1 mM EDTA
0.5 M $NaHPO_4$, pH 7.2 (134 g $Na_2HPO_4 \cdot 7H_2O$ plus 4 ml 85% H_3PO_4/liter = 1 M $NaHPO_4$)
7% SDS

Low-stringency wash buffer I

2× SSC (*APPENDIX 2*)
0.1% SDS

Low-stringency wash buffer II

0.5% BSA (fraction V)
1 mM Na_2EDTA
40 mM $NaHPO_4$, pH 7.2
5% SDS

Sonicated herring sperm DNA, 2 mg/ml

Resuspend 1 g herring sperm DNA (Boehringer Mannheim #223636) in a convenient volume (about 50 ml of water) by sonicating briefly. The DNA is now ready to be sheared into short molecules by sonication. Place the tube containing the herring sperm DNA solution in an ice bath (the tube must be stable even if the ice begins to melt). The sonicator probe is placed in the DNA solution (without touching the bottom of the vessel). The sonicator is turned on to 50% power 20 min, or until there is a uniform and obvious decrease in viscosity. At no time should the tube containing the DNA become hot to the touch. After sonication, the DNA is diluted

to a final concentration of 2 mg/ml, frozen in 50-ml aliquots, and thawed as needed.

COMMENTARY

Background Information

All hybridization methods depend upon the ability of denatured DNA to reanneal when complementary strands are present in an environment near but below their T_m (melting temperature). In a hybridization reaction involving double-stranded DNA on a filter and a single-stranded DNA probe there are three different annealing reactions occurring. First, there are the desired probe-DNA interactions which result in signal. Second, there are mismatch interactions that occur between related but nonhomologous sequences; these mismatch hybrids are the ones that must be eliminated during the washing of the filters. Non-sequence-specific interactions also occur and these result in noise. The ability to extract information from a particular hybridization experiment is a function of the signal-to-noise ratio. High background or poor specific signal can both result in uninterpretable results.

Washing nitrocellulose filters is required to remove excess radioactive probe, as well as radioactive probe that has bound to the DNA on the filter as mismatch hybrids. Temperature and salt concentration dramatically affect the maintenance of specific hybrids. Detergents and other charged species can have a profound effect upon the nonspecific binding of species that contribute to background. In this protocol, hybridization is achieved in a solution containing 50% formamide. Excess probe is rinsed away under low-stringency conditions so that further hybridization will not occur. Once the hybridization solution is rinsed away, it is possible to proceed to a high-stringency wash without fear of further hybridization. When washing is complete, the filters should produce very little "noise" when monitored with a Geiger counter. Although single-copy sequence probe normally does not produce a signal that is detectable with a Geiger counter, a probe corresponding to more abundant sequences will produce a signal that can be "heard" with a Geiger counter.

Literature Review

Hybridization to filter membranes forms a basis of recombinant DNA technology and is described in detail earlier in the manual (*UNIT 2.9*). The protocols described here vary from those used for Southern blot filter hybridization in that the volume of the hybridization is usually larger and the washing conditions are different. Dextran sulfate is an important component of the hybridization solution as it increases the rate of reassociation of the nucleic acids.

The protocols in this unit describe methods for hybridizing radioactive probes to membrane-bound plaques or colonies. These procedures for screening recombinant clones were first suggested by Grunstein and Hognes (1975) and by Benton and Davis (1977). The conditions of hybridization proposed in the basic protocol involving hybridization in formamide was originally described by Denhardt (1966) and Gillespie and Spiegelman (1965) while the alternate protocol using aqueous hybridization solution was introduced by Church and Gilbert (1984).

The method of washing filters under stringent conditions to remove background was first proposed by Southern (1975). Botchan et al. (1976) described the benefit of adding SDS to the wash solution. Jeffreys and Flavell (1977) first employed the wash conditions described in the protocols presented here.

Critical Parameters

Hybridization. Kinetically, the hybridization of DNA (or RNA) probes to filter-bound DNA is not significantly different from hybridization in solution. For single-stranded probes, the rate of hybridization follows first-order kinetics, since probe is available in excess. Under conditions of excess probe, the time for hybridization is inversely proportional to the probe concentration. For double-stranded probes the rate of hybridization displays a more complex relationship to the initial probe concentration. However, to a first approximation the initial probe concentration is inversely proportional to the rate of hybridization. To determine the actual time required for the successful hybridization of a given probe, either empirical data must be available or the following formula can be used to determine the length of time (in hours) required to achieve 50% hybridization (T_{50}):

$$\tfrac{1}{2} \times \tfrac{x}{5} \times \tfrac{y}{10} \times 2 = T_{50}$$

where x is the weight of probe in micrograms; y is the complexity of probe in kilobases; and z is the volume of hybridization solution in milliliters. The length of time T is given in hours. Maximum hybridization signal will be obtained if the reaction is allowed to proceed to $5 \times T_{50}$, although 1 to $2 \times T_{50}$ is often used.

It is also clear that nonspecific interactions

occur and that in any hybridization, sources of noise will be present. Therefore, from a practical standpoint one conventionally utilizes concentrations of nick-translated probe on the order of 1 to 15 ng/ml of hybridization, where the specific activity of the probe is from 5×10^7 cpm/μg to $>10^8$ cpm/μg. Too much probe in a hybridization is as bad as too little.

One important source of background hybridization to filters is due to the hybridization of the probe to vector sequences or to *E. coli* DNA. Be certain that there is no vector or *E.coli* DNA sequences in the probe. This can best be ensured by isolating the probe from one type of vector (e.g., plasmid) and screening a library made with a different type of vector (e.g., bacteriophage).

Washing temperature. Washing at low stringency is a straightforward proposition. Buffer is added at room temperature and washing proceeds at room temperature.

High-stringency wash is determined empirically. The relative homology between the probe and target sequence is a determining parameter. If the homology is 100%, a high temperature (65° to 75°C) can be used. As the homology drops, lower washing temperatures must be used. In general one starts at 37° to 40°C, raising the temperature by 3° to 5°C intervals until background is low enough not to be a major factor in the autoradiography.

The length of the probe is also important. Very short probes (<100 bp) must be washed at lower temperatures, even if the homology is 100%. Washing strategy is the same as for probes of differing homology.

Salt concentration. The lower the salt concentration, the higher the stringency. With this said, the protocols as outlined do not require adjustment of salt concentration for adjustment of stringency. Only the washing temperature is varied.

Probe. The nucleic acid probe must be of high specific activity and greater than 50 bp in length so that it can form stable hybrids.

Anticipated Results

After washing the filters the background should be barely detectable with a Geiger counter.

With a high-specific-activity probe $>5 \times 10^7$ cpm/μg and an overnight hybridization reaction with a 1-kb unique sequence probe, hybridizing bacterial colonies or bacteriophage plaques can be visualized after a 1 to 18 hr exposure.

Time Considerations

Generally hybridizations are carried on overnight for 12 to 16 hr. This is sufficient for most probes and blots. However, with probes of increasing complexity longer hybridization times are required. This is preferable to increasing the probe concentration from experiment to experiment.

Autoradiography requires 1 to 18 hr.

Literature Cited

Benton, W.D. and Davis, R.W. 1977. Screening λgt recombinant clones by hybridization to single plaques in situ. *Science* 196:180.

Botchan, M., Topp, W., and Sambrook, J. 1976. The arrangement of simian virus 40 sequences in the DNA of transformed cells. *Cell* 9:269-287.

Church, G. and Gilbert, W. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. U.S.A.* 81:1991-1995.

Denhardt, D. 1966. A membrane filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* 23:641-646.

Gillespie, D. and Spiegelman, S. 1965. A quantitative assay for DNA-RNA hybrids with DNA immobilized on a membrane. *J. Mol. Biol.*12:829-842.

Grunstein, M. and Hogness, D. 1975. Colony Hybridization: A method for the isolating of cloned DNA's that contain a specific gene. *Proc. Natl. Acad. Sci. U.S.A.* 72:3961.

Jeffreys, A.J. and Flavell, R.J. 1977. A physical map of the DNA region flanking the rabbit β globin gene. *Cell* 12:429-439.

Southern, E.M. 1975. Detection of specific sequence among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98:503-517.

Contributed by William M. Strauss
Harvard Medical School
Boston, Massachusetts